Optimizing a Ranking Function of Web Search Agents

Optimization Lab.
최성준

Question
Why the global optimization algorithm is needed to set an optimal parameter of ranking function?

Contents
- Motivation
- Three elements needed for optimizing
  - Training set of queries
  - Ranking function model
  - Performance function
- Optimization method
- Pitfalls
  - Overfitting problem
  - Expressiveness of ranking function model
- Conclusion

Motivation
- Ranking Function
  - Orders documents based on their estimated relevance to the user’s query
  - Have a large number of free parameters
    - The weights on terms in documents
    - Parameters of the similarity metric
    - Retrieval threshold
  - Can have a significant effect on the performance of a system (web search agent, more generally, retrieval system)

Three elements needed to optimize
1. Training set of queries
2. Ranking function model
3. Performance function that estimate the specific ranking function’s performance

Objective
We want to find an optimal set of parameters of ranking function model that maximizes the performance function value with respect to training set of queries
Training set of queries

- Needed to train a ranking function
- Generally given in forms of pairs ([6])
  (query, relevant documents)

Ranking function model

- Decide the approximate structure of ranking functions
- Can limit the domain and the search space of functions ([2], [4])

---

Ranking function model

- Example ([2])
  \[ R_{\Theta, q}(d) = \frac{q \cdot d}{\sum d_i} \]
  - \( q \): query vector derived from sample \( q \)
  - \( d \): document vector from sample \( d' \)
  - \( \Theta = (\Theta_1, \Theta_2) \)
  - If \( \Theta_2 = 0 \), \( R_{\Theta, q}(d) \) is just inner product

---

Performance function

- Standard
  - Average precision for training set of queries ([1], [2], [5])
  - Average precision is hard to handle
  - Alternative
    - Designed to manipulate it easily
    - Must be highly correlated to average precision

Optimization method I

- Conjugate Gradient ([2])
  - Gradient-based
  - Intuitive and greedy method
  - Performance function must be differentiable
Optimization method I
- Conjugate Gradient
  - Average precision is not differentiable
  - Alternative performance function is needed to use conjugate gradient method
  - Example [12]

\[
J(R_{n,k}(d)) = \frac{1}{|A|} \sum_{d\in A} \frac{1}{\sum_{d'\in A} R_{n,k}(d')}
\]

Optimization method I
- Conjugate Gradient
  - Limit
    - Can fall into local optima

Optimization method II
- Simulated Annealing [7]
  - Stochastic approximation method
  - Probabilistic hill-climbing strategy
  - Performance function don't have to be differentiable

Optimization method II
- Simulated Annealing
  - Global optimization algorithm [3], [4]
Pitfalls

- Overfitting problem
  - Use hold-out set ([13])

- Expressiveness of ranking function model
  - The model must be rich enough to cover a wide search space
  - After choosing such a rich model, optimization method is used to set an optimal parameter of the model

Conclusion

- The effect of selection of specific ranking function
  - Can have a significant effect on the performance
- Global optimization algorithm
  - Ranking function search space ([3], [4], [7])
- Pitfalls
  - Overfitting problem
  - Expressiveness of ranking function model

References
